Justice Quarterly

ISSN: 0741-8825 (Print) 1745-9109 (Online) Journal homepage: https://www.tandfonline.com/loi/rjqy20

What Exactly Is the Bargain? The Sensitivity of Plea Discount Estimates

Shi Yan

To cite this article: Shi Yan (2020): What Exactly Is the Bargain? The Sensitivity of Plea Discount Estimates, Justice Quarterly, DOI: 10.1080/07418825.2019.1707856

To link to this article: https://doi.org/10.1080/07418825.2019.1707856

Published online: 03 Jan 2020.

Submit your article to this journal

Article views: 73

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rjqy20

JUSTICE QUARTERLY
https://doi.org/10.1080/07418825.2019.1707856

Check for updates

What Exactly Is the Bargain? The Sensitivity of Plea Discount Estimates

Shi Yan

School of Criminology and Criminal Justice, Arizona State University, Phoenix, AZ, USA

ABSTRACT

Defendants who plead guilty receive less harsh sentences than those convicted at trial. Theories suggest that the magnitude of the plea discount reflects local courtroom norms, and also guides the behavior of defendants. The testing of theories on plea bargaining relies on credible estimates of the size of the plea discount. This study found that plea discount estimates were sensitive to the dependent variable used. The data used included the full sample of felony defendants in New York State who pled guilty, in which nearly half were not incarcerated. When a severity score that encompassed the severity of probation was used instead of the length of incarceration, the plea discount estimate decreased by 26 percentage points for the full sample. This pattern was found in all subsamples as well.

ARTICLE HISTORY
Received 26 March 2019
Accepted 17 December 2019

KEYWORDS
Plea bargaining; severity scale; alternative sanctions; counterfactual

Introduction

The vast majority of criminal cases in the USA are disposed of with guilty pleas (Reaves, 2013). The prevalence of guilty pleas has led to the argument that plea bargaining has factually changed the foundation of American criminal justice (e.g. Church, 1979; Dervan, 2019; Lynch, 2016; Stuntz, 2011). Over the decades, researchers have paid extensive attention to many aspects of plea bargaining, such as the trends in plea bargaining (Shaeffer, 2019), defendants' voluntariness of pleading decisions (Hussemann & Siegel, 2019), and the "value" of plea decisions (King, Soulé, Steen, & Weidner, 2005; Ulmer & Bradley, 2006). With the infusion of theories and methods from multiple disciplines (Abrams, 2011; Redlich, Bibas, Edkins, & Madon, 2017; Yan, 2020), these works have considerably strengthened our understanding of guilty pleas.

Despite the progress, one common practice in plea bargaining research is to focus exclusively on the incarceration sentences – either the "in or out" outcome or the length of incarceration (Baumer, 2013; Bibas, 2004). This overemphasis on incarceration is inconsistent with theories and practices of plea bargaining. While the high stakes

CONTACT Shi Yan shiyan@asu.edu
This data is provided by the New York State Division of Criminal Justice Services (DCJS). The opinions, findings, and conclusions expressed in this publication are those of the author and not those of DCJS. Neither New York State nor DCJS assumes liability for its contents or use thereof.
© 2019 Academy of Criminal Justice Sciences

associated with incarceration warrant extensive research, alternative sanctions are also integral parts of "a rational sentencing system" (Morris & Tonry, 1990).[1] Alternative sanctions are particularly relevant in plea bargaining research, as they serve as effective incentives for prosecutors to obtain guilty pleas (Alschuler, 1968; Heumann, 1978). Moreover, as this paper will note, the neglect of alternative sanctions in empirical research may result in systematic overestimation of the disparity between trial and plea sentences.

Recently, courts and sentencing researchers have begun to shift away from the so-called modal approach (Baumer, 2013) by taking a much broader range of decision points into consideration. A group of recent empirical studies adopted the "cumulative advantage" framework, contending that pretrial decisions, such as bail, dismissal, and charge reduction, helped shaping the sentence (Johnson & Larroulet, 2019; Kohler-Hausmann, 2018; Kutateladze, Andiloro, Johnson, & Spohn, 2014; Metcalfe & Chiricos, 2018; Scott-Hayward & Fradella, 2019). Another group of studies explored the various aspects of alternative sanctions (e.g. Martin, Sykes, Shannon, Edwards, & Harris, 2018; Robinson, 2018). The present study adds to this ongoing effort at the intersection of the two focal points. Using case processing data from New York State, this paper demonstrates how the measurement of alternative sanctions affects the estimates of plea discounts. Specifically, the plea discount estimates for the entire sample decreased by more than 40% (26 percentage points) after considering the severity of probation. This pattern was also found in all the subsamples analyzed.

Making sense of the plea discount

The Sixth Amendment guarantees the defendant's right to a jury trial. Nevertheless, most defendants today plead guilty instead of taking advantage of this right. From an organizational perspective, plea bargaining serves the goals of the courtroom workgroup – disposing of the caseload, maintaining group cohesion, and "doing justice" by rewarding defendants who express remorse and assume responsibility (Church, 1979; Eisenstein & Jacob, 1977; Heumann, 1978). Under a more recent and refined theoretical framework, defendants who plead guilty warrant less harsh sentences because of the focal concerns the courtroom actors face: defendants' blameworthiness, protection of the community, and practical constraints (Ulmer & Bradley, 2006; Ulmer, Eisenstein, & Johnson, 2010). Prosecutors and judges may view defendants who plead guilty as less blameworthy, less risky, and more cooperative in addressing the courts' practical constraints. These theories suggest that the magnitude of plea discounts (and any variation therein) reflects local courtroom norms and "going rates" (Nardulli, Eisenstein, & Flemming, 1988; Ulmer, 1997).

Meanwhile, critics of plea bargaining claim that large disparities between the sentences at trial and plea, often phrased and perceived as "the trial penalty," are potentially coercive (Andersson & Robinson, 2019; Lynch, 2016; National Association of Criminal Defense Lawyers, 2018). Many studies reported anecdotal stories of

[1]For simplicity, the term "alternative sanctions" refers to non-incarceration sentences throughout the paper. As presented later in the paper, the use of this term does not necessarily designate incarceration as the default sentencing option.

defendants who received considerably harsher sentences after declining the plea offers (Caldwell, 2011; Lynch, 2016; Rakoff, 2014). Since the overall probability of conviction at trial is high (Bushway & Redlich, 2012; Reinganum, 1988), such large disparities increase the stakes of going to trial and push rational defendants towards guilty pleas (Zottoli, Daftary-Kapur, Winters, & Hogan, 2016). In the most extreme cases, the pressure can result in innocent defendants pleading guilty (Cooper, Meterko, & Gadtaula, 2019; Ochoa & Salazar, 2019; Redlich & Shteynberg, 2016; Schneider & Zottoli, 2019). Stuntz (2011) has commented as follows,

> The lack of careful investigation that characterizes most felony prosecutions virtually guarantees that a significant number of innocent defendants are pressured to plead to crimes they did not commit. And within the much larger universe of guilty defendants, those who are punished most severely are often *those who made the worst deals*, not *those who committed the worst crimes* [emphasis added]. (p. 58)

The theory "bargaining in the shadow of trial," originating from the law and economics research, provides the best illustration of the consequence of large plea discounts/trial penalties (hereinafter the shadow model, Landes, 1971; Mnookin & Kornhauser, 1979).[2] The classic representation of the theory is the equation below,

$$s_P = p * s_T,\ \ 0 < p < 1,\ \ s_P \geq 0,\ \ s_T \geq 0 \tag{1}$$

where s_P stands for the sentence at plea, and s_T stands for the sentence at trial. The parameter p stands for the probability of conviction at trial, which is typically associated with the strength of evidence. If s_P is lower than $p * s_T$, then rational defendants are better off pleading guilty – and vice versa.

According to the National Association of Criminal Defense Lawyers (2018, pp. 20–21), for defendants processed in the federal system, the average sentence imposed on those who pled guilty was 3.3 years, and the average trial sentence was 10.8 years. Putting those numbers into Equation (1), it suggests that a defendant would be better off pleading guilty as long as the probability of conviction at trial was 30.6% or higher. The average trial sentence was 10 times higher than the plea sentence for some crime types, meaning that nearly every defendant – or every defendant with a 10% chance of conviction at trial – should plead guilty.

The paragraphs above highlight the crucial role of the size of the plea discount or the trial penalty. Under the courtroom workgroup and focal concerns perspectives, the size of the plea discount is a main indicator of local courtroom norms and dynamics (Nardulli et al., 1988). Under the shadow model, the relative magnitude of the plea discount directly guides the rational behavior of defendants (Bushway et al., 2014). Under either theoretical framework, systematic overestimation or underestimation of the plea discounts can severely distort the understanding of the system. The question is – how accurate are the estimates in the extant literature?

[2]For a thorough review and test of the shadow model, see Bushway, Redlich, and Norris (2014). This study has no intention to validate the theory, but only utilizes the framework to depict why large plea discounts are problematic. For details on challenges to the shadow model, see Bibas (2004), Redlich, Bibas, et al. (2017), and Yan and Bushway (2018).

The role of alternative sanctions in plea bargaining

A common practice in quantitative research on plea bargaining is to use the incarceration outcome – either the in/out decision or the length of incarceration – as the only dependent variable in the analyses (known as "the modal approach," Baumer, 2013). This practice treats those not incarcerated as if they received zero punishment, which is inconsistent with either of the theoretical frameworks presented above. Specifically, under the courtroom workgroup and focal concerns framework, alternative sanctions are different from no punishment. When prosecutors and judges believe that punishment as unnecessary, the most common practice is a dismissal of the case. By contrast, alternative sanctions often carry a non-negligible punitive component (Alschuler, 1968; Heumann, 1978; Kimchi, 2019; Wodahl, Ogle, Kadleck, & Gerow, 2013). Under the shadow model, the two sides of Equation (1) are *the severities of the sentences*, not necessarily *the length of incarceration*. If all alternative sanctions are assumed to have zero or near-zero harshness, those who plead guilty to an alternative sanction will appear as if they received a 100% plea discount. This can result in a systematic overestimation of the plea discount.

Moreover, researchers have found strong support for the severity of alternative sanctions, as they often lead to non-negligible time, monetary, and social costs (often known as collateral consequences). For instance, individuals on probation often face restrictions on their daily activities, as well as the risk of revocation and re-incarceration (Kimchi, 2019; Kohler-Hausmann, 2018; Watts, 2016). Monetary sanctions – often perceived as the most lenient form of formal sanctioning – can impose significant difficulty for economically disadvantaged individuals (Harris, 2016; Martin et al., 2018; Scott-Hayward & Fradella, 2019). Multiple studies found that defendants perceived some alternative sanctions to be harsher than short incarceration terms, and may prefer to serve a shorter incarceration term than a longer probation term (Crouch, 1993; May & Wood, 2010; Petersilia & Deschenes, 1994; Wodahl et al., 2013). All these findings stand in stark contrast to the practice of treating alternative sanctions as zero punishments. Moreover, statistics show that alternative sanctions are far more frequent than incarceration sentences (Harris, 2016; Kaeble, 2018). By modeling only incarceration sentences, researchers neglect the more substantial portion of plea bargaining outcomes.

Although the theories support the importance of alternative sanctions, the empirical specification is much more difficult. To date, researchers have yet to reach a consensus on the standard practice of modeling alternative sanctions. One group of studies focused on modeling the qualitative difference among the sanctions and focused on the combinations of restrictions and sanction types (Engen, Gainey, Crutchfield, & Weis, 2003; Johnson & DiPietro, 2012; Kimchi, 2019; Yan, 2017). Another group of studies attempted to generate unified scales for different types of sanctions (Leclerc & Tremblay, 2016; May & Wood, 2010; Pina-Sánchez et al., 2019; Spelman, 1995; Wodahl et al., 2013). Unlike the former group, these studies emphasized the exchange rates among the sentence types (e.g. how many months of probation are equal to one month of prison). Both approaches have advantages and disadvantages, and neither has become the sole dominant approach in empirical studies (Johnson, King, & Spohn, 2016; Ulmer, 2012). Given the research question, the present study will

adopt the second approach and compare plea discount estimates using a severity scale to those resulting from the use of length of incarceration.

Method

Data and samples

The present study used the Computerized Criminal History (CCH) data of New York State, maintained by and obtained from the New York State Division of Criminal Justice Services (DCJS). The CCH dataset includes the population of fingerprintable felony and misdemeanor arrests made under the jurisdiction of New York State, which guarantees the sample size. It also contains detailed information on the type and the seriousness of the top charges at arraignment and conviction. The full analytic sample consisted of all defendants arrested in New York State between 2008 and 2012, and who were arraigned for a felony and pled guilty ($n = 348{,}232$). Additionally, the counterfactual models were based on all felony defendants convicted at a trial during the same observation period ($n = 7{,}019$).

The empirical models and the dependent variables

The counterfactual regression models

Smith (1986) introduced an approach of estimating counterfactual trial sentences for defendants who pled guilty. To date, researchers still regard this approach as "the best previous empirical work" (Abrams, 2011, p. 203) and have widely adopted it in empirical studies (Johnson & Larroulet, 2019; Metcalfe & Chiricos, 2018; Yan & Bushway, 2018). This approach starts with a regression model explaining the sentence of defendants convicted at trial,

$$\begin{aligned} \textit{Sentence at trial} = \ & \beta_0 + \beta_1 \ \textit{current crime} + \ \beta_2 \ \textit{criminal history} \\ & + \beta_3 \textit{extralegal variables} + e \end{aligned} \tag{2}$$

Then for defendants who pled guilty, it is possible to predict the counterfactual trial sentence using the coefficients in Equation (2) (i.e. the sentence they would have received at trial; for additional technical details of the empirical method, see references in the paragraph above).

The dependent variables

This study adopted two different variables to specify the sentences in Equations (1) and (2). The first one, commonly used in the literature, was *the length of post-sentence incarceration*, measured in months (top coded at 480 months for those received life sentences and sentences longer than 480 months).[3] There was no differentiation between prison and jail lengths. The second variable was a *sentence severity score*, generated as a combined score from three types of sentences: prison, jail, and

[3]There was no special coding for life sentences without parole since only 1% among those sentenced to life sentences (and less than 0.05% of the total sample) had the "without parole" clause.

Table 1. Severity scores for select sanctions.

	Length				Parameters	
Sanction type	6 months	12 months	24 months	36 months	α	β
Prison						
Spelman's original work	–	119	–	159		
Scale in the current study	–	117.46	140.65	156.29	61.56	0.26
Jail						
Spelman's original work	–	100	–	–		
Scale in the current study	69.38	100.18	–	–	26.84	0.53
Probation						
Spelman's original work	35	37	51	–		
Scale in the current study	31.91	41.24	53.3	61.92	16.44	0.037

Note: Spelman' original work contained other terms not presented here, for full details, see Spelman (1995, p. 120). See Equation (3) for the role of parameters α and β.

probation.[4] This score is adopted from the work of Spelman (1995) and Yan (2017). Spelman surveyed a group of individuals serving different sentences and generated a severity score for a variety of specific prison, jail, and probation terms, as presented in Table 1.

For the severity of each sanction type (prison, jail, and probation), Spelman (1995) gave only a few point estimates as presented above. To accommodate any given sentence length, Yan (2017) converted these point estimates into a set of scales under two assumptions. First, for any sanction type, a length of zero corresponds to a zero severity score (for more on the validity of this assumption, see Methodological Notes below). Second, for any given sanction type, the relationship between the length and the harshness is not necessarily linear. Psychological and behavioral economic research suggests that sentences are also subject to diminishing returns. Therefore, as the length of the sentence increases, the perceived harshness associated with the additional length will decrease (Crouch, 1993; Tversky & Kahneman, 1974). The assumptions lead to the following power functional form for each sanction type,

$$S_{sanction} = \alpha Q^{\beta}_{sanction}\ ,\ \ sanction = prison,\ \ jail,\ \ probation \tag{3}$$

where Q stands for the length of each sanction, and S stands for the corresponding harshness. The parameters α and β are obtained from an OLS regression logging both sides.[5]

The independent variables

All regression models included a set of legal and extralegal variables to explain and predict the sentences. *The seriousness of the top arraignment charge* had five categories, class A to class E felony (class A being the most serious and class E being the least serious). *Crime type* was coded into 14 categories, based primarily on the categories used in the State Court Processing Statistics (Reaves, 2013) with some modification considering their prevalence and importance in the sample (see Results section for the categories). The models also included *the counts of the top arraignment charge* and *the numbers of prior felony and misdemeanor convictions*. For the extralegal variables, the models included the defendants' *race* (coded as White, Black, and other/

[4]The data-granting agency, DCJS, advised that the coding of probation lengths in the CCH dataset was more reliable and accurate than some other types of alternative sanctions.

[5]Logging both sides of Equation (3) gets $lnS = ln\alpha + \beta lnQ$.

missing), *sex*, *ethnicity* (coded as Hispanic and non-Hispanic), *age*, and *age squared* to account for the non-linear relationship between age and the sentence. The models also controlled for the *county of disposition* and *year* to account for unobserved factors related to them. All regression models clustered standard errors by county.

Methodological notes

Three pieces of clarification are necessary here on the methodology stated above. First, one criticism of the analytic approach (Equation 2) is that it does not create genuine counterfactuals. Specifically, defendants who went to trial may have different unobserved characteristics from defendants who pled (such as risk preference, legal representation, and cognitive traits, see Bibas, 2004; Redlich, Bibas, et al., 2017). If that was the case, the predicted trial counterfactuals might be subsequently biased. Although this concern is valid, there is no feasible way to correct for this potential bias, as the correction requires data beyond what most researchers can obtain (Baumer, 2013). There are some efforts in the literature to address this problem, but these alternative strategies have shortcomings too.[6] To date, Smith's (1986) modeling strategy is still frequently used in investigating the size of plea discounts (e.g. Bushway & Redlich, 2012; Johnson & Larroulet, 2019; Metcalfe & Chiricos, 2018).

Second, there are also potential concerns regarding the external validity of the severity scale, given both the strong parametric assumption (Equation 3 above, see also Yan, 2017) and that Spelman's (1995) study was conducted in a different research site (Texas) more than two decades ago. Although the challenges are legitimate, the primary purpose of the current study is to *reveal* the dependent variable problem rather than to *fix* it. Therefore, there is no intention to claim that the scale used here is the only possible measure of the sentence severities. Several other groups of researchers have also attempted to measure the relative severity of sanctions (Crouch, 1993; May & Wood, 2010; Petersilia & Deschenes, 1994), and researchers today are still developing sanction severity scales from exchange rates (Leclerc & Tremblay, 2016; Pina-Sánchez et al., 2019; Wodahl et al., 2013).[7] To some extent, all these scales are potentially useful (or not useful) for the present study. This study used Spelman's scale because other studies either contained an insufficient amount of information to develop scales appropriate for the current study (Crouch, 1993; Petersilia & Deschenes, 1994; Wodahl et al., 2013) or led to unreasonable severity score estimates (May & Wood, 2010).[8]

[6]Specifically, Abrams (2011) adopted an instrumental variable approach. He used the judges' length of tenure as an instrument to estimate the effect of pleading on sentence length. While the method had its merits, it was also built on strong assumptions, and the instrument was relatively weak (p. 217). Another possible way to measure some of the unobserved variables (such as cognitive traits and risk preference) is through experimental design. A group of studies explored the role of psychological factors in guilty pleas (e.g. Dervan & Edkins, 2013; Redlich & Shteynberg, 2016; Tor, Gazal-Ayal, & Garcia, 2010). Yet the focus was mostly on the decision to plead guilty rather than the size of plea discounts. These studies also have known external validity issues.

[7]Although both scales by Leclerc and Tremblay and Pina-Sánchez were more recent, they were developed outside of the United States (one in England and Wales and the other in Canada). Although Wodahl et al.'s scale was developed in the United States, it did not contain the exchange rates between prison and other types of sentences.

[8]Specifically, when I developed May and Wood's point estimates into scales following the procedure described above, the severity of long probation terms (3 years or longer) well exceeded the severity of prison terms of the same length. This issue is likely to be due to the parametric assumption in Equation (3). To the contrary, as shown in Table 1, the severity scores generated from Spelman's study were more reasonable. In either case, I have no intention to challenge the measures of the original studies, since neither group of researchers appeared to intend to develop the point estimates into scales and to provide out-of-bound estimates.

Third, both dependent variables used in this study – the incarceration length and the severity score – were built on *post-conviction* sentences. The dataset used did not contain information on pretrial detention, which has garnered considerable attention among both researchers and policymakers (Kohler-Hausmann, 2018; Scott-Hayward & Fradella, 2019). For defendants, many outcomes before the sentencing stage are punitive, and this is particularly true for those detained prior to trial (Feeley, 1979a, 1979b; Lowenkamp, VanNostrand, & Holsinger, 2013). Without information on pretrial detention, the analysis can suffer from a systematic underestimation of the severity of sentences. For an individual defendant, such a downward bias may be substantial, particularly when the sentence imposed is time served. However, recent statistics suggest that on average, the length of pretrial detention is much shorter than that of post-conviction sentences (Chauhan et al., 2017; Leslie & Pope, 2017).[9] Although the lack of pretrial detention information is not ideal, it unlikely to have substantial consequences on a large sample – the full sample of felony defendants who pled guilty in a state in a five-year period.

Results

Descriptive statistics

Table 2 presents the descriptive statistics for the samples. Among defendants who pled guilty, 52.6% received some form of incarceration, and the average incarceration length of these defendants was 10.8 months (20.6 months for those incarcerated). Given that all defendants started with a felony charge, the use of incarceration in this sample was much less frequent than the national average, thus highlighting the measurement issue of the dependent variable (Reaves, 2013). Over 60% of defendants were arraigned for a class D or E felony, and the most frequent arraignment crime types were burglary, larceny, drug trafficking, and drug possession (each over 10%). About 40% of defendants were Black, and slightly over a third were Hispanic. The sample was predominantly male (85%) and had an average age of 31.6 years.

Table 2 also presents the descriptive statistics for defendants convicted at trial. Approximately 85% of these defendants were incarcerated, and the average sentence length was 117.5 months (138.2 months for those incarcerated). Even though both the probability and the length of incarceration were considerably higher, this does not necessarily indicate the size of the plea discount given the difference in observed characteristics between the two defendant groups. Compared with the plea sample, defendants convicted at trial faced more serious charges and were more likely to face violent charges. They were also more likely to be Black, slightly older, and had more prior felony convictions.

[9]Chauhan et al. (2017, p. 38) reported that for felony defendants in New York City, the median length of pretrial detention was around 10 days, and the average length was around 75 days. In the current sample, the average sentence for New York City defendants who pled was 12.44 months – nearly five times as long as the average length of pretrial detention. In the full plea sample, fewer than 5% of defendants received a term of time served.

The trial sentence models

Table 3 presents the regression models explaining the sentence received by defendants convicted at trial for the full sample (Equation 2). For both dependent variables – incarceration length and the severity score – nearly all legal variables included were significant predictors of the sentence, although the coefficients differed. Defendants charged with class B to E felonies received much shorter sentences than those charged with class A felonies. Except for those charged with sex crimes, defendants charged with other crime types received shorter sentences than those charged with murder. Each additional count of the arraignment charge, as well as each prior felony conviction, was associated with longer sentences, whereas the number of prior misdemeanor convictions had a small and negative association with both dependent variables. In both models, male defendants received significantly harsher sentences than females, and Black defendants received significantly harsher sentences than Whites.

The R^2 values suggest that the models explained over 60% of the variance in sentence lengths and over 50% of the variance in the severity scores. Although the models did not include some essential variables such as the strength of evidence and legal representation, the R^2 values are still high (Spohn, 2015; Ulmer, 2012). Given the absence of presumptive and restrictive sentencing guidelines in New York, the fit of the models appears to be good.[10]

Full sample plea discount estimates

Main point estimates

With the coefficients from these models, the next step is to predict the counterfactual trial sentence for each defendant in the sample (Johnson & Larroulet, 2019; Metcalfe & Chiricos, 2018). Table 4 presents the results for the entire sample on the top. The average sentence length for the plea sample was 10.8 months. If these defendants were convicted at trial, they would instead have received an average sentence of 29.0 months. The average plea discount for the full sample was therefore 62.6%. However, the average severity scores at the plea and trial for the plea sample were 74.5 and 117.5 respectively, which led to an average plea discount of 36.6%. With the severity of probation taken into account, the observed gap between the trial and plea sentences was reduced by 26 percentage points (or 41.5%).

Differentiation of two possible explanations

There are two possible explanations for the observed difference between the two plea discount estimates. The first possibility is that the use of the severity score reduced the number of defendants who had a value of zero in the plea sentence and therefore reduced the prevalence of "100% plea discounts" in the sample. Of the defendants who pled guilty, 47.4% had an incarceration length of zero, whereas

[10]The Penal Law of New York State prescribes very large range of discretion at the sentencing stage. For example, with no predicate records, a defendant convicted of a class B violent felony can be sentenced to a determinate term between 5 and 25 years in prison. A defendant convicted of a class B non-violent-non-drug felony can be sentenced to an indeterminate term with the upper bound ranging between 3 and 25 years. This stands in sheer contrast with jurisdictions that have presumptive sentencing guidelines and narrowly prescribed sentence intervals, such as the federal jurisdiction (Lynch, 2016).

Table 2. Descriptive statistics.

	(1)		(2)	
	Plea sample		Trial conviction sample	
	M	*SD*	*M*	*SD*
Incarceration	0.53	0.50	0.85	0.36
Incarceration length	10.84	28.44	117.57	156.27
Sentence severity score	74.46	57.11	167.08	86.18
Class A felony (e.g. murder first degree, criminal possession of a controlled substance first degree)	0.01	0.12	0.14	0.34
Class B felony (e.g. assault first degree, robbery first degree, grand larceny first degree)	0.21	0.41	0.31	0.46
Class C felony (e.g. robbery second degree, grand larceny second degree, criminal possession of a controlled substance fourth degree)	0.14	0.35	0.21	0.41
Class D felony (e.g. assault second degree, grand larceny third degree, criminal possession of a controlled substance fifth degree)	0.33	0.47	0.22	0.41
Class E felony (e.g. grand larceny fourth degree, criminal mischief third degree)	0.31	0.46	0.13	0.34
Murder	0.01	0.08	0.13	0.34
Sex crimes	0.03	0.16	0.09	0.29
Robbery	0.09	0.29	0.13	0.33
Aggravated assault	0.07	0.26	0.15	0.36
Simple assault	0.05	0.23	0.04	0.20
Burglary	0.10	0.30	0.10	0.30
Larceny	0.13	0.34	0.06	0.23
Fraud/forgery	0.06	0.24	0.02	0.13
Criminal mischief/arson	0.03	0.18	0.02	0.14
Drug trafficking	0.11	0.31	0.05	0.21
Drug possession	0.12	0.32	0.05	0.22
Weapons crimes	0.04	0.20	0.08	0.27
Driving crimes	0.08	0.27	0.05	0.22
Nuisance	0.07	0.26	0.04	0.19
Arraignment charge counts	1.20	9.40	1.28	1.85
Prior felony convictions	0.70	1.30	1.03	1.49
Prior misdemeanor convictions	2.81	6.39	2.78	6.29
White	0.53	0.50	0.38	0.49
Black	0.44	0.50	0.59	0.49
Other or missing race	0.03	0.17	0.03	0.17
Hispanic	0.36	0.48	0.32	0.47
Male	0.85	0.36	0.91	0.28
Age	31.62	11.98	33.42	11.41
n	348,232		7019	

only 23.3% had a severity score of zero.[11] The second possibility is that the severity scale converted the relationship between the sentence length and the severity into a non-linear functional form with diminishing returns. In most cases, the ratio between the two sentence lengths is larger than the ratio between the corresponding severity scores. To differentiate the two mechanisms, Table 4 presents the

[11]Out of defendants who had a severity score of zero, 66.2% received conditional discharge, whereas 23.3% received time served. I was unable to model the severity of either sanction types due to the lack of comparable measures in the literature (conditional discharge) and missing information (time served, see Methodological Notes above).

Table 3. Regression models explaining the sentence.

	(1) Length of incarceration	(2) Sentence severity score
Class B felony	−216.32***	−72.28***
	(7.08)	(3.01)
Class C felony	−271.90***	−106.95***
	(8.40)	(4.41)
Class D felony	−323.09***	−149.96***
	(7.99)	(5.07)
Class E felony	−322.36***	−158.69***
	(7.69)	(4.21)
Sex Crimes	−12.36	3.02
	(7.74)	(3.23)
Robbery	−28.60*	−9.62*
	(10.83)	(4.48)
Aggravated assault	−40.84***	−19.37***
	(9.00)	(4.63)
Simple assault	−62.57***	−40.40***
	(9.49)	(5.47)
Burglary	−42.18***	−25.99***
	(10.68)	(5.92)
Larceny	−70.93***	−30.99***
	(8.53)	(3.72)
Fraud/forgery	−59.32***	−36.17***
	(11.52)	(5.81)
Criminal mischief/arson	−62.66***	−44.93***
	(8.70)	(4.78)
Drug trafficking	−147.89***	−68.59***
	(10.85)	(5.39)
Drug possession	−151.40***	−78.41***
	(9.24)	(3.64)
Weapons crimes	−44.24***	−15.71*
	(8.35)	(6.24)
Driving crimes	−65.38***	−38.87***
	(9.19)	(6.11)
Nuisance	−70.45***	−30.38***
	(7.85)	(4.81)
Arraignment charge counts	1.40***	1.14***
	(0.38)	(0.27)
Prior felony convictions	22.81***	11.91***
	(2.87)	(1.09)
Prior misdemeanor convictions	−1.30***	−0.27
	(0.23)	(0.25)
Black	4.78*	4.27**
	(1.88)	(1.41)
Other race	7.48	4.89
	(4.41)	(3.83)
Hispanic	−1.65	−0.52
	(2.62)	(2.02)
Male	16.68***	20.17***
	(3.50)	(1.76)
Age	3.06**	1.61**
	(1.00)	(0.59)
Age squared	−0.04**	−0.02**
	(0.01)	(0.01)
Constant	307.50***	242.60***
	(16.73)	(10.52)
n	7019	7019
R^2	0.63	0.56

Note: Robust standard errors in parentheses. Year and county fixed effects controlled.

***$p < 0.001$.

**$p < 0.01$.

*$p < 0.05$.

Table 4. Plea discount estimates for full sample and subsamples.

	(1) Value/Estimate
Length of incarceration	
Average sentence at plea	10.84
Average sentence at trial	29.02
Estimated plea discount	62.64%
Sentence severity score	
Average sentence at plea	74.46
Average sentence at trial	117.50
Estimated plea discount	36.63%
Severity score for incarceration	
Average sentence at plea	53.05
Average sentence at trial	100.11
Estimated plea discount	47.01%

results using *the severity score for incarceration only*. This dependent variable was the step between the two used in the main analysis. It preserved the non-linear relationship between the sentence length and the severity but also treated all non-incarceration sentences as zeroes. The estimated plea discount was 47.0% for the entire sample, falling in between the two main estimates, slightly closer to the estimate obtained using the severity score. This finding suggests that both factors – the reduction of zeroes and the change of functional form – appeared to contribute to the discrepancy between the plea discount estimates obtained using the incarceration length and the severity score.

Distribution of individual plea discounts

All analyses so far have only presented point estimates (averages) for the entire sample. At the individual level, each defendant also had a predicted plea discount. Figures 1 and 2 present the distributions of the individual-level plea discounts estimates from the predicted plea sentences and the actual plea sentences, respectively.[12] At the individual level, the counterfactual approach can lead to unreasonable estimates of the plea discount – some individuals had a negative plea discount (or a "plea penalty," Abrams, 2011), whereas others had a plea discount over 100%.[13] Since some of the estimates were extreme and unlikely to be realistic (much lower than −100% or much higher than 100%), all histograms used only cases in the middle 90% of the distribution (i.e. for each given plea discount estimate, those cases between the 5th and the 95th percentiles). Both figures demonstrate that the use of sentence severity scores led to much smaller plea discount estimates at the individual level. Moreover,

[12]Both estimates compare the plea sentence against the predicted counterfactual trial sentence. The common practice to estimate the individual-level plea discount is to use *the predicted plea sentence* (estimated from Equation 2, using the plea sentence as the dependent variable on the plea sample) rather than *the actual plea sentence*, as the former accounts for the systematic components in sentencing better (Johnson & Larroulet, 2019; Piehl & Bushway, 2007). The current study presents both because the estimates with the actual plea sentence better demonstrate the consequence of zero sentences (or "100% plea discounts").

[13]A negative plea discount occurs when the estimated trial sentence is positive but lower than the plea sentence, whereas a plea discount over 100% occurs when the estimated trial sentence is negative. These could either be merely random noise or bear a substantive meaning (for a theory on "the plea penalty," see Abrams, 2011). It is also noteworthy that much fewer individuals had an unreasonable plea discount estimate when the severity score served as the dependent variable (Figure 1). Nevertheless, the present study simply leaves the values as they are since the interpretation of these unreasonable values is beyond the scope of the current study.

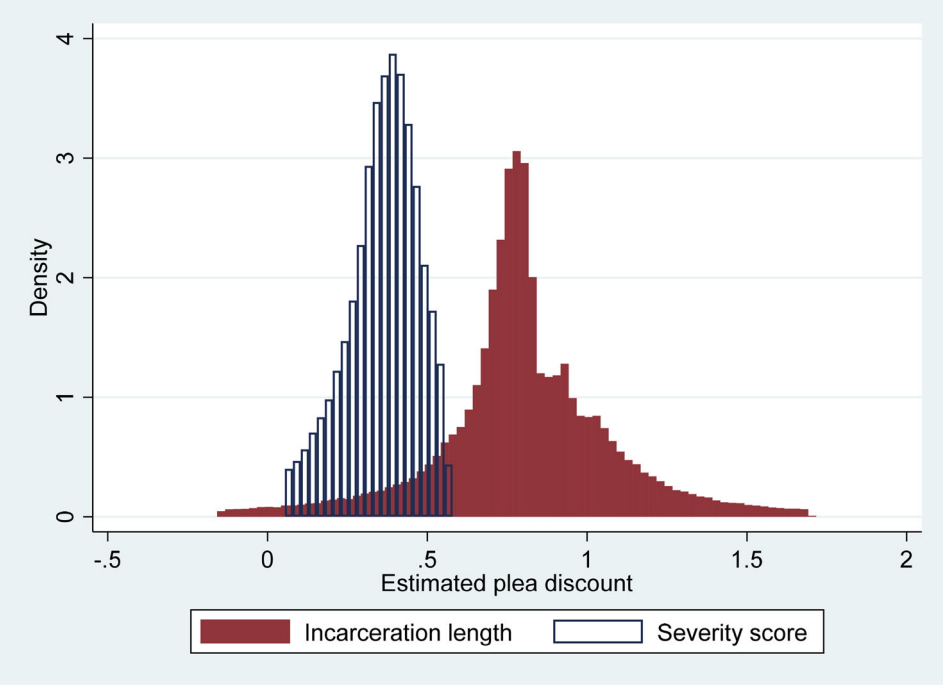

Figure 1. Histogram of estimated individual plea discounts using predicted plea sentence and predicted trial sentence.
Note: For both plea discount estimates, only those between the 5th and the 95th percentiles used.

Figure 2 suggests that substantially fewer individuals received an estimated 100% plea discount when the severity score dependent variable was used than when incarceration length was used.

Subsample analyses

Given the level of disagreement between the two plea discount estimates, it is necessary to investigate whether the plea discount is more sensitive to the dependent variable for some defendants than for others. The next step was a series of analyses on different subsamples of plea defendants. Each of the subsample models used only the corresponding subset of defendants convicted at trial (e.g. for the subsample for New York City, only defendants convicted at trial in New York City were used to estimate Equation 2). Most of the subsamples were based on the seriousness and the type of the arraignment charge. Although a certain level of correlation existed between the two variables, the correlation was not problematically high because the Penal Law set different degrees of charges for most of the common crime types.[14] In all those models, there did not appear to be a multicollinearity problem (VIFs for all crime seriousness and crime type variables were smaller than 10). For all results presented below,

[14]For example, for both robbery and burglary, the first degree was a class B felony, the second degree was a class C felony, and the third degree was a class D felony. For both criminal sale of a controlled substance and criminal possession of a controlled substance, the first to the fifth degree were distributed through classes A to D.

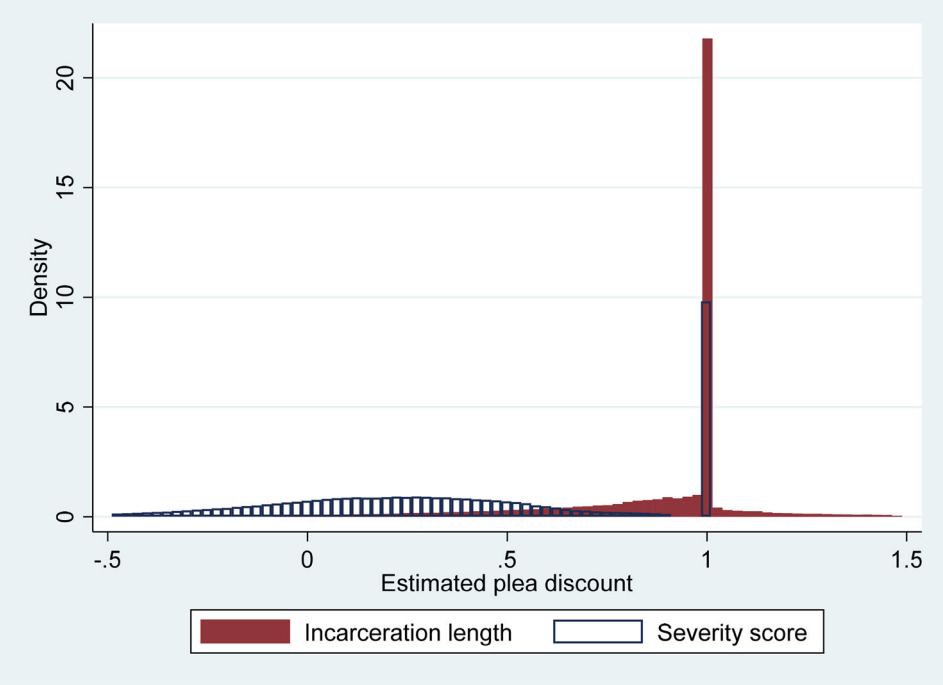

Figure 2. Histogram of estimated individual plea discounts using actual plea sentence and predicted trial sentence.
Note: For both plea discount estimates, only those between the 5th and the 95th percentiles used.

the comparison is primarily between the two dependent variables on each subsample rather than across the different subsamples.

Table 5 presents the findings for defendants processed in New York City ($n = 138{,}391$), defendant subsamples based on the seriousness of the current crime (classes A and B, classes C and D, and class E, n ranging between 78,160 and 163,479), and subsamples based on the general type of the current crime (violent, property, drug, public order, n ranging between 67,378 and 115,411). In all these subsamples, there was a noticeable difference between the plea discount estimates using incarceration length and using the sentence severity score. For example, for defendants in New York City, the average plea discount estimated using incarceration length was 62.7%, and the average plea discount estimated using the severity score was 42.0% – a 20.7 percentage point disparity. The same pattern was also the case for all other subsamples, with disparities in the plea discount estimates ranging between 22.8 and 37.0 percentage points. It did not appear that any certain subsample was driving the main findings.

Plea discount estimates from the Tobit model?

Other than the analytic strategies presented so far, one alternative approach in sentencing research is to use the Tobit model to account for the heterogeneity among non-incarceration sentences (Tobin, 1958). The Tobit model treats the incarceration length as a censored dependent variable, and at one time was common among quantitative studies of sentencing (Albonetti, 1997; Piehl & Bushway, 2007). As a cautionary

Table 5. Plea discount estimates for subsamples.

	(1) Length of incarceration	(2) Sentence severity score
New York City subsample ($n = 138,391$)		
Average sentence at plea	12.44	72.10
Average sentence at trial	33.33	124.38
Estimated plea discount	62.68%	42.03%
Classes A and B subsample ($n = 78,160$)		
Average sentence at plea	23.63	93.57
Average sentence at trial	66.19	159.81
Estimated plea discount	64.29%	41.45%
Classes C and D subsample ($n = 163,479$)		
Average sentence at plea	9.5	74.83
Average sentence at trial	29.54	117.39
Estimated plea discount	67.84%	36.25%
Classes E subsample ($n = 106,593$)		
Average sentence at plea	3.53	59.87
Average sentence at trial	10.43	91.97
Estimated plea discount	66.19%	34.91%
Violent subsample ($n = 87,123$)		
Average sentence at plea	18.47	87.78
Average sentence at trial	54.86	133.68
Estimated plea discount	66.33%	34.34%
Property subsample ($n = 115,411$)		
Average sentence at plea	7.29	65.03
Average sentence at trial	20.65	104.29
Estimated plea discount	64.70%	37.64%
Drug subsample ($n = 78,320$)		
Average sentence at plea	10.65	72.78
Average sentence at trial	41.28	135.38
Estimated plea discount	74.20%	46.24%
Public order subsample ($n = 67,378$)		
Average sentence at plea	7.30	75.33
Average sentence at trial	22.69	108.99
Estimated plea discount	67.83%	30.88%

analysis, I attempted to replicate the analysis above with a Tobit model rather than an OLS model. Table 6 presents the finding on the full sample as well as the subsamples based on the type of the current crime, analogous to those presented in Tables 4 and 5. In three of the five analyses, the average plea discount was undefined because the estimated plea sentence was negative. In the other two analyses, the estimated average plea discounts were huge – both well over 90%. The reason for the findings is that the sheer number of zero incarceration terms shifted the entire sentence length distribution towards the left by a large margin under the Tobit specification.[15] There is no intention to nullify the potential use of Tobit and other censored dependent variable models in sentencing research, particularly when the goal is to estimate the coefficient for a specific variable. However, the results here attest that the Tobit model

[15]A key difference here is that when the researcher uses the Tobit model to conduct the counterfactual analysis, both plea and trial sentences need to be predicted values under the Tobit model (for details on the technique, see Piehl & Bushway, 2007, pp. 111–113). By definition, on a defendant sample who pled, the average predicted sentence at plea from an OLS model is equal to the observed average plea sentence. However, this is not the case for the Tobit model. Most defendants in Piehl and Bushway's study were incarcerated, and their plea discount estimates were still reasonable despite the Tobit model shifted the distributions to the left as well (Piehl & Bushway, 2007; Tables 1 and 4). In the present study, the average plea sentences for all samples in Table 6 were all considerably lower than the corresponding actual average plea sentence because of the prevalence of the zeros.

Table 6. Plea discount estimates using the Tobit model on length of incarceration for full and select subsamples.

	(1) Length of incarceration
Full sample ($n = 348,232$)	
Average sentence at plea	−5.28
Average sentence at trial	1.97
Estimated plea discount	–
Violent subsample ($n = 87,123$)	
Average sentence at plea	1.58
Average sentence at trial	34.95
Estimated plea discount	95.5%
Property subsample ($n = 115,411$)	
Average sentence at plea	−9.93
Average sentence at trial	−18.71
Estimated plea discount	–
Drug subsample ($n = 78,320$)	
Average sentence at plea	1.17
Average sentence at trial	30.97
Estimated plea discount	96.2%
Public order subsample ($n = 67,378$)	
Average sentence at plea	−5.10
Average sentence at trial	−1.76
Estimated plea discount	–

Note: Estimated plea discount undefined if either the plea sentence or the trial sentence is below zero.

may not be an appropriate tool when the goal is prediction of the sentence and when non-incarceration sentences are common in the sample.

Discussion

The prevalence of plea bargaining has led to pronounced changes to the court system today (Dervan, 2019; Lynch, 2016; Stuntz, 2011). In their recent review of social scientific research on plea bargaining, Redlich, Wilford, and Bushway (2017, pp. 467–468) summarized the current research on plea bargaining as follows,

> On the one hand, many questions remain about who pleads guilty, why, and under what circumstances. Although there are theoretical models of plea decision making, they remain underdeveloped and undertested … On the other hand, to be sure, there has been a renewed interest in guilty pleas by scholars from several different disciplines [citations omitted]. There have also been recent rulings by the U.S.'s highest court toward regulating the plea process, prompting Justice Scalia to opine that "a whole new boutique of constitutional jurisprudence ('plea bargaining law')" was created ….

A more fundamental question must be answered before the theoretical and empirical advances in the area of "plea bargaining law" can take place: what exactly is the value of the plea? Numerous works, both qualitative and quantitative, suggest that the parties negotiate around *the sentence*, which may, but does not always include, *incarceration* (Heumann, 1978; May & Wood, 2010; Nardulli et al., 1988; Ulmer, 1997). The current controversy on plea bargaining, particularly the notion of "the trial penalty," is primarily built on the sheer disparity between the sentences offered to those who plead guilty and the threatened sentences at trial (Lynch, 2016; National Association of Criminal Defense Lawyers, 2018; Rakoff, 2014). Theorists from the courtroom workgroup and focal concern perspectives would contend that the size of plea

discounts reflects local courtroom norms, as well as the assessment of blameworthiness and risk (Ulmer & Bradley, 2006; Ulmer et al., 2010). Theorists from the shadow model would contend that the size of plea discounts guide defendants' behavior (Bushway et al., 2014; Zottoli et al., 2016). Despite the different intellectual roots, neither theory precludes alternative sanctions from the core arguments and aims at solely explaining the disparities in incarceration. To the contrary, both have underscored the vital role of alternative sanctions.

Empirical tests of all these theories rely on the size of plea discount estimates. Instead of validating specific theories, the present study contributes to the literature by revealing the sensitivity of the estimates to the measurement of the sentence, which I hope will serve as a cautionary note for future researchers. The study found that the estimates depend on whether the severity of probation is included. When the models only included the sentence length, the sample had an estimated plea discount of 62.6%. However, the average plea discount dropped to 36.6% when the severity score, which included the probation sentences (Spelman, 1995; Yan, 2017), served as the dependent variable. The estimates of the disparity between plea and trial sentences became much more modest when probation was taken into consideration. The analysis further found that two factors contributed simultaneously to the discrepancy: the reduction of "zero sentences" and the consideration of diminishing returns for the sentences.

Substantively, the empirical findings call for a balanced view on the plea bargaining process and the outcome. On the one hand, the disparity between the sentences at trial and at plea is indeed concerning (National Association of Criminal Defense Lawyers, 2018; Stuntz, 2011). On the other hand, the actual sentence defendants receive at plea is also worth attention. Many alternative sanctions – including but not limited to probation – are still considerably punitive on defendants (Crouch, 1993; Petersilia & Deschenes, 1994; Wodahl et al., 2013). The level of harshness associated with a plea will be even higher if the subsequent criminal records and collateral consequences are taken into account (Denver, Pickett, & Bushway, 2018; Kohler-Hausmann, 2018). In the plea bargaining context, what the defendants *actually received* has received much less attention than what they *could have received at trial*. One can never overemphasize the significance of the latter in the light of the Sixth Amendment. However, we also need to be aware that the former is what actually affects the defendants' lives.

It is necessary to emphasize that the purpose of this study is not to validate different plea discount estimates against some known, proven value. Given the highly complicated nature of the calculation of the costs and pains of punishment, the true severity of the sentence may be specific to each defendant (Bibas, 2004; Ostrom, Ostrom, & Kleiman, 2004). To *some* defendants, the harshness of alternative sanctions could be negligible, and an alternative sanction plea offer may actually be a 100% discount (Alschuler, 1968; Heumann, 1978). However, there are reasons to believe that this may not be the case for *all* defendants. The purpose of the study is not to evaluate the existence and degree of coercion in plea bargaining either. Despite the policy prominence of this issue (Lynch, 2016; Rakoff, 2014), it requires much more information than what this dataset (and perhaps this type of datasets) can provide.

This study also found that the sensitivity of plea discount estimates appeared to exist regardless of the case characteristics. In all subsample analyses, there was a

visible gap between the two estimates, ranging between 21 to nearly 40 percentage points. In other words, *no specific group of defendants, regardless of charge type or charge seriousness, was immune to this problem*. Moreover, the empirical analysis also revealed that the Tobit model – a once-popular and reasonable approach in modeling sentence severity – may not be appropriate at least for this particular research question and this analytic sample. Even though the Tobit model could potentially help to correct for the selection bias in the slope estimates, its parametric assumption was strong and inconsistent with prior research on alternative sanctions. This finding echoes recent empirical work that recommends researcher to move away from Tobit models (Wilson, Loughran, & Brame, 2018), and highlights the necessity of the continued development of modeling strategies.

Other than the issues already discussed in detail in the Method section, this study has some other limitations as well. First, this dataset did not have the information on the full sets of charges for each defendant. It only had the top charge (type and counts) for each case and did not have less serious charges filed. For example, a defendant charged of armed robbery may face one count of robbery and one count of unlawful possession of a firearm. The bargaining around the less serious charges is potentially important, but cannot be examined using the current data. Second, this dataset does not have some key variables that potentially determine the sentence, such as the type of legal counsel and the strength of evidence. Therefore, there may be bias in the plea discount estimates due to these missing variables. Third, the main findings – the instability of plea discount estimates between dependent variables – may be particular to this sample. In jurisdictions and places where the use of incarceration is common, the plea discount estimates may be much less sensitive to the dependent variable. Although a variety of academic and governmental publications have suggested the wide variation in the use of incarceration versus alternative sanctions among counties and states (Reaves, 2013; Watts, 2016), much more work is necessary to evaluate the extent of the problem at other jurisdictions.

Despite the limitations, the present study moves the current research on plea bargaining one step forward. It provides a cautionary note on the operationalization of the plea discount or the trial penalty, since the way researchers quantify the discounts can directly affect the policy and theoretical implications. Future research should continue to explore possible ways to capture the variation in sentences – not only among the incarceration terms, but also among the combinations of alternative sanctions. Studies that employ mixed methods and utilize multi-jurisdiction data will be necessary to answer these important research questions regarding plea bargaining.

Acknowledgements

I thank Shawn Bushway, Hank Fradella, Cassia Spohn, Gary Sweeten, and the three anonymous reviewers for their comments. I also thank Jason Walker for the copyediting.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Shi Yan is an assistant professor at the School of Criminology and Criminal Justice, Arizona State University. His research focuses on plea bargaining, sentencing, and measurement issues in crime and criminal justice research.

ORCID

Shi Yan http://orcid.org/0000-0002-6484-7687

References

Abrams, D. S. (2011). Is pleading really a bargain? *Journal of Empirical Legal Studies*, *8*(s1), 200–221. doi:10.1111/j.1740-1461.2011.01234.x

Albonetti, C. A. (1997). Sentencing under the federal sentencing guidelines: Effects of defendant characteristics, guilty pleas, and departures on sentence outcomes for drug offenses, 1991–1992. *Law & Society Review*, *31*(4), 789–822. doi:10.2307/3053987

Alschuler, A. W. (1968). The prosecutor's role in plea bargaining. *The University of Chicago Law Review*, *36*(1), 50–112. doi:10.2307/1598832

Andersson, E., & Robinson, J. (2019). The insidious injustice of the trial penalty: "It is not the intensity but the duration of pain that breaks the will to resist". *Federal Sentencing Reporter*, *31*(4–5), 222–225. doi:10.1525/fsr.2019.31.4-5.222

Baumer, E. P. (2013). Reassessing and redirecting research on race and sentencing. *Justice Quarterly*, *30*(2), 231–261. doi:10.1080/07418825.2012.682602

Bibas, S. (2004). Plea bargaining outside the shadow of trial. *Harvard Law Review*, *117*(8), 2463–2547. doi:10.2307/4093404

Bushway, S. D., & Redlich, A. D. (2012). Is plea bargaining in the "shadow of the trial" a mirage? *Journal of Quantitative Criminology*, *28*(3), 437–454. doi:10.1007/s10940-011-9147-5

Bushway, S. D., Redlich, A. D., & Norris, R. J. (2014). An explicit test of plea bargaining in the "shadow of the trial". *Criminology*, *52*(4), 723–754. doi:10.1111/1745-9125.12054

Caldwell, H. M. (2011). Coercive plea bargaining: The unrecognized scourge of the justice system. *Catholic University Law Review*, *61*(1), 63–96.

Chauhan, P., Hood, Q. O., Balazon, E. M., Cuevas, C., Lu, O., Tomascak, S., & Fera, A. G. (2017). *Trends in custody: New York City Department of Correction, 2000–2015*. New York, NY: John Jay College of Criminal Justice.

Church, T. W., Jr. (1979). In defense of "bargain justice". *Law & Society Review*, *13*(2), 509–525. doi:10.2307/3053266

Cooper, G. S., Meterko, V., & Gadtaula, P. (2019). Innocents who plead guilty: An analysis of patterns in DNA exoneration cases. *Federal Sentencing Reporter*, *31*(4–5), 234–238. doi:10.1525/fsr.2019.31.4-5.234

Crouch, B. M. (1993). Is incarceration really worse? Analysis of offenders' preferences for prison over probation. *Justice Quarterly*, *10*(1), 67–88. doi:10.1080/07418829300091711

Denver, M., Pickett, J. T., & Bushway, S. D. (2018). Criminal records and employment: A survey of experiences and attitudes in the United States. *Justice Quarterly*, *35*(4), 584–613. doi:10.1080/07418825.2017.1340502

Dervan, L. E. (2019). Arriving at a system of pleas: The history and state of plea bargaining. In V. A. Edkins & A. D. Redlich (Eds.), *A system of pleas: Social science's contributions to the real legal system* (pp. 11–23). New York, NY: Oxford University Press.

Dervan, L. E., & Edkins, V. A. (2013). The innocent defendant's dilemma: An innovative empirical study of plea bargaining's innocence problem. *The Journal of Criminal Law and Criminology*, *103*(1), 1–48.

Eisenstein, J., & Jacob, H. (1977). *Felony justice: An organizational analysis of criminal courts*. Boston, MA: Little, Brown.

Engen, R. L., Gainey, R. R., Crutchfield, R. D., & Weis, J. G. (2003). Discretion and disparity under sentencing guidelines: The role of departures and structured sentencing alternatives. *Criminology*, *41*(1), 99–130. doi:10.1111/j.1745-9125.2003.tb00983.x

Feeley, M. M. (1979a). Pleading guilty in lower courts. *Law & Society Review*, *13*(2), 461–466. doi:10.2307/3053263

Feeley, M. M. (1979b). *The process is the punishment: Handling cases in a lower criminal court*. New York, NY: Russell Sage Foundation.

Harris, A. (2016). *A pound of flesh: Monetary sanctions as punishment for the poor*. New York, NY: Russell Sage Foundation.

Heumann, M. (1978). *Plea bargaining: The experiences of prosecutors, judges, and defense attorneys*. Chicago, IL: The University of Chicago Press.

Hussemann, J., & Siegel, J. (2019). Pleading guilty: Indigent defendant perceptions of the plea process. *Tennessee Journal of Law and Policy*, *13*(2), 459–522.

Johnson, B. D., & DiPietro, S. M. (2012). The power of diversion: Intermediate sanctions and sentencing disparity under presumptive guidelines. *Criminology*, *50*(3), 811–850. doi:10.1111/j.1745-9125.2012.00279.x

Johnson, B. D., King, R. D., & Spohn, C. (2016). Sociolegal approaches to the study of guilty pleas and prosecution. *Annual Review of Law and Social Science*, *12*(1), 479–495. doi:10.1146/annurev-lawsocsci-110615-084755

Johnson, B. D., & Larroulet, P. (2019). The "distance traveled": Investigating the downstream consequences of charge reductions for disparities in incarceration. *Justice Quarterly*, *36*(7), 1229-1257. doi:10.1080/07418825.2018.1529250

Kaeble, D. (2018). *Probation and* parole *in the United States, 2016*. Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice. Retrieved from https://www.bjs.gov/content/pub/pdf/ppus16.pdf

Kimchi, A. (2019). Investigating the assignment of probation conditions: Heterogeneity and the role of race and ethnicity. *Journal of Quantitative Criminology*, *35*(4), 715–745. doi:10.1007/s10940-018-9400-2

King, N. J., Soulé, D. A., Steen, S., & Weidner, R. R. (2005). When process affects punishment: Differences in sentences after guilty plea, bench trial, and jury trial in five guidelines states. *Columbia Law Review*, *105*(4), 959–1009.

Kohler-Hausmann, I. (2018). *Misdemeanorland: Criminal courts and social control in an age of broken windows policing*. Princeton, NJ: Princeton University Press.

Kutateladze, B. L., Andiloro, N. R., Johnson, B. D., & Spohn, C. (2014). Cumulative disadvantage: Examining racial and ethnic disparity in prosecution and sentencing. *Criminology*, *52*(3), 514–551. doi:10.1111/1745-9125.12047

Landes, W. M. (1971). An economic analysis of the courts. *The Journal of Law and Economics*, *14*(1), 61–107. doi:10.1086/466704

Leclerc, C., & Tremblay, P. (2016). Looking at penalty scales: How judicial actors and the general public judge penal severity. *Canadian Journal of Criminology and Criminal Justice*, *58*(3), 354–384. doi:10.3138/cjccj.2014.E31

Leslie, E., & Pope, N. G. (2017). The unintended impact of pretrial detention on case outcomes: Evidence from New York City arraignments. *The Journal of Law and Economics*, *60*(3), 529–557. doi:10.1086/695285

Lowenkamp, C. T., VanNostrand, M., & Holsinger, A. (2013). *The hidden costs of pretrial detention*. Houston, TX: Laura and John Arnold Foundation.

Lynch, M. (2016). *Hard bargains: The coercive power of drug laws in federal court*. New York, NY: Russell Sage Foundation.

Martin, K. D., Sykes, B. L., Shannon, S., Edwards, F., & Harris, A. (2018). Monetary sanctions: Legal financial obligations in US systems of justice. *Annual Review of Criminology*, *1*(1), 471–495. doi:10.1146/annurev-criminol-032317-091915

May, D. C., & Wood, P. B. (2010). *Ranking correctional punishments: Views from offenders, practitioners, and the public*. Durham, NC: Carolina Academic Press.

Metcalfe, C., & Chiricos, T. (2018). Race, plea, and charge reduction: An assessment of racial disparities in the plea process. *Justice Quarterly*, *35*(2), 223–253. doi:10.1080/07418825.2017.1304564

Mnookin, R. H., & Kornhauser, L. (1979). Bargaining in the shadow of the law: The case of divorce. *The Yale Law Journal*, *88*(5), 950–997. doi:10.2307/795824

Morris, N., & Tonry, M. (1990). *Between prison and probation: Intermediate punishments in a rational sentencing system*. New York, NY: Oxford University Press.

Nardulli, P. F., Eisenstein, J., & Flemming, R. B. (1988). *The tenor of justice: Criminal courts and the guilty plea process*. Urbana, IL: University of Illinois Press.

National Association of Criminal Defense Lawyers. (2018). *The trial penalty: The Sixth Amendment right to trial on the verge of extinction and how to save it*. Washington, DC: National Association of Criminal Defense Lawyers.

Ochoa, C., & Salazar, C. (2019). How the threat of the trial penalty coerces the innocent to plead guilty: A first-hand account of an exoneree. *Federal Sentencing Reporter*, *31*(4–5), 299–302. doi:10.1525/fsr.2019.31.4-5.299

Ostrom, C. W., Ostrom, B. J., & Kleiman, M. (2004). *Judges and discrimination: Assessing the theory and practice of criminal sentencing*. Washington, DC: National Institute of Justice.

Petersilia, J., & Deschenes, E. P. (1994). Perceptions of punishment: Inmates and staff rank the severity of prison versus intermediate sanctions. *The Prison Journal*, *74*(3), 306–328. doi:10.1177/0032855594074003003

Piehl, A. M., & Bushway, S. D. (2007). Measuring and explaining charge bargaining. *Journal of Quantitative Criminology*, *23*(2), 105–125. doi:10.1007/s10940-006-9023-x

Pina-Sánchez, J., Gosling, J. P., Chung, H.-I., Bourgeois, E., Geneletti, S., & Marder, I. D. (2019). Have the England and Wales guidelines affected sentencing severity? An empirical analysis using a scale of severity and time-series analyses. *The British Journal of Criminology*, *59*(4), 979–1001. doi:10.1093/bjc/azz005

Rakoff, J. S. (2014, November 20). Why innocent people plead guilty. *The New York Review of Books*. Retrieved from http://www.nybooks.com/articles/2014/11/20/why-innocent-people-plead-guilty/

Reaves, B. A. (2013). Felony defendants in large urban counties, 2009 – Statistical tables. Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice. Retrieved from https://www.bjs.gov/content/pub/pdf/fdluc09.pdf

Redlich, A. D., Bibas, S., Edkins, V. A., & Madon, S. (2017). The psychology of defendant plea decision making. *American Psychologist*, *72*(4), 339–352. doi:10.1037/a0040436

Redlich, A. D., & Shteynberg, R. V. (2016). To plead or not to plead: A comparison of juvenile and adult true and false plea decisions. *Law and Human Behavior*, *40*(6), 611–625. doi:10.1037/lhb0000205

Redlich, A. D., Wilford, M. M., & Bushway, S. (2017). Understanding guilty pleas through the lens of social science. *Psychology, Public Policy, and Law*, *23*(4), 458–471. doi:10.1037/law0000142

Reinganum, J. F. (1988). Plea bargaining and prosecutorial discretion. *The American Economic Review*, *78*(4), 713–728.

Robinson, G. (2018). Delivering McJustice? The probation factory at the Magistrates' court. *Criminology & Criminal Justice*, *19*(5), 605–621. doi:10.1177/1748895818786997

Schneider, R. A., & Zottoli, T. M. (2019). Disentangling the effects of plea discount and potential trial sentence on decisions to plead guilty. *Legal and Criminological Psychology*, *24*(2), 288–304. doi:10.1111/lcrp.12157

Scott-Hayward, C. S., & Fradella, H. F. (2019). *Punishing poverty: How bail and pretrial detention fuel inequalities in the criminal justice system*. Oakland, CA: University of California Press.

Shaeffer, R. (2019). The trial penalty: An international perspective. *Federal Sentencing Reporter*, *31*(4–5), 321–330. doi:10.1525/fsr.2019.31.4-5.321

Smith, D. A. (1986). The plea bargaining controversy. *The Journal of Criminal Law and Criminology*, *77*(3), 949–968. doi:10.2307/1143445

Spelman, W. (1995). The severity of intermediate sanctions. *Journal of Research in Crime and Delinquency*, *32*(2), 107–135. doi:10.1177/0022427895032002001

Spohn, C. (2015). Race, crime, and punishment in the twentieth and twenty-first centuries. *Crime and Justice*, *44*(1), 49–97. doi:10.1086/681550

Stuntz, W. J. (2011). *The collapse of American criminal justice*. Cambridge, MA: The Belknap Press of Harvard University Press.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, *26*(1), 24–36. doi:10.2307/1907382

Tor, A., Gazal-Ayal, O., & Garcia, S. M. (2010). Fairness and the willingness to accept plea bargain offers. *Journal of Empirical Legal Studies*, *7*(1), 97–116. doi:10.1111/j.1740-1461.2009.01171.x

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. doi:10.1126/science.185.4157.1124

Ulmer, J. T. (1997). *Social worlds of sentencing: Court communities under sentencing guidelines*. Albany, NY: SUNY Press.

Ulmer, J. T. (2012). Recent developments and new directions in sentencing research. *Justice Quarterly*, *29*(1), 1–40. doi:10.1080/07418825.2011.624115

Ulmer, J. T., & Bradley, M. S. (2006). Variation in trial penalties among serious violent offenses. *Criminology*, *44*(3), 631–670. doi:10.1111/j.1745-9125.2006.00059.x

Ulmer, J. T., Eisenstein, J., & Johnson, B. D. (2010). Trial penalties in federal sentencing: Extra-guidelines factors and district variation. *Justice Quarterly*, *27*(4), 560–592. doi:10.1080/07418820902998063

Watts, A. (2016). *Probation in-depth: The length of probation sentences*. Minneapolis, MN: Robina Institute of Criminal Law and Criminal Justice.

Wilson, T., Loughran, T., & Brame, R. (2018). Substantial bias in the Tobit estimator: Making a case for alternatives. *Justice Quarterly*. Advance online publication. 1–27. doi:10.1080/07418825.2018.1517220

Wodahl, E. J., Ogle, R., Kadleck, C., & Gerow, K. (2013). Offender perceptions of graduated sanctions. *Crime & Delinquency*, *59*(8), 1185–1210. doi:10.1177/0011128709333725

Yan, S. (2017). Search for the hidden punishments: An alternative approach to studying alternative sanctions. *Journal of Quantitative Criminology*, *33*(1), 21–44. doi:10.1007/s10940-015-9275-4

Yan, S. (2020). Estimating the size of plea discounts: Why does it matter? In C. Spohn & P. K. Brennan (Eds.), *Handbook on sentencing policies and practices in the 21st century* (pp. 188–207). New York, NY: Routledge.

Yan, S., & Bushway, S. D. (2018). Plea discounts or trial penalties? Making sense of the trial-plea sentence disparities. *Justice Quarterly*, *35*(7), 1226–1249. doi:10.1080/07418825.2018.1552715

Zottoli, T. M., Daftary-Kapur, T., Winters, G. M., & Hogan, C. (2016). Plea discounts, time pressures, and false-guilty pleas in youth and adults who pleaded guilty to felonies in New York City. *Psychology, Public Policy, and Law*, *22*(3), 250–259. doi:10.1037/law0000095